Vagueness Facilitates Search

Kees van Deemter

University of Aberdeen
k.vdeemter@abdn.ac.uk,
WWW home page: http://www.csd.abdn.ac.uk/∼kvdeemte/

Abstract. This paper addresses the question why language is vague. A novel answer to this question is proposed, which complements other answers suggested in the literature. It claims that vagueness can facilitate search, particularly in quasi-continuous domains (such as physical size, colour, or temperature), given that different speakers are likely to attach subtly different meanings to words (such as "tall", "blue", or "hot") defined over such domains.

1 Introduction

Two questions dominate theoretical research on vagueness. The first is of a logical-semantic nature: *What formal models offer the best understanding of vagueness?* Many answers to this question have been proposed (e.g. [1], [2] for an overview), but none of these has found general acceptance so far. The second question is of a pragmatic nature and asks *Why is language vague?* This question has been asked forcefully by the economist Barton Lipman, who has shown that some seemingly plausible answers resist analysis in terms of classical Game Theory [3], [4]. While a number of tentative answers to this question have been suggested (for a survey, see [6], [7]), Lipman's question is still partly unresolved, particularly with respect to situations where there is no conflict between the speaker and the hearer (cf. [8]).

The present paper will focus on the second question, and in doing so it will obtain some insights into the first question as well. We will elaborate on a novel answer to this question, which was sketched in broad outline in [6], [7], explaining the probabilistic basis of the argument, and discussing what we see as its merits more fully than before. In a nutshell, we argue primarily that vagueness can facilitate search. Additionally, we argue that Partial Logic is better placed to explain this phenomenon than Classical Logic, and that theories that give pride of place to *degrees* (including many-valued logics [9], but also two-valued theories that include degrees, e.g., [10]) are even better placed than Partial Logic to do this. We do not claim that facilitation of search is the only rationale for vagueness, or that degrees are necessary for explaining the benefits of vagueness: a non-quantitative model involving an ordinal scale might be equally suitable.

2 Informal outline of the argument

Let's call a domain *quasi-continuous* if it contains objects which resemble each other so much that they are indistinguishable. Domains do not have to be *mathematically* continuous to have this property: it suffices for them to contain objects that are similar enough in the relevant dimension (a person of 180.1cm and one of 180.2cm height, for example) that they cannot be told apart given the measurement tools at hand. Examples abound, including the heights of all the people you know, or all the colours that you have seen.

In a quasi-continuus domain, it is difficult for people to align the meanings of the predicates defined over them: there are bound to be people that one speaker calls 'tall' that another does not. The causes include physical and cultural differences between people. David Hilbert, for example, who focusses on colour terms, explains how the differences in people's eyes (e.g., in terms of the density of pigment layers on the lens and the retina, in in terms of the sensitivity of the photo receptors) make it unavoidable that one normally sighted person can often distinguish between colour patches where another cannot [11]. The role of cultural issues was highlighted in [12], where it was shown that different weather forcasters use different criteria in their use of temporal phrases; according to some forecasters, for example, the start of the *evening* has something to do with dinner time, whereas for others, the time when the sun sets is more relevant, while yet others believe that the time on the clock is the only relevant consideration. Rohit Parikh has written insightfully about such matters, and we shall use and adapt one of his examples to present our own argument below.[1]

In Parikh's original story of Ann and Bob, Ann asks Bob to find her book on topology, adding that "it is blue" [13]. Ann and Bob use different concepts of 'blue', but if the overlap between them, as compared to their symmetric difference, is sufficiently large then Ann's utterance may still be very useful, because it may reduce the time that Bob should expect to take before finding the referent. All the same, the mismatch between speaker and hearer does cause Bob's search for the topology book to take more time than it would otherwise have done. This is particularly true because the book, b, may be an element of $\|blue\|_{Ann}\| - \|blue\|_{Bob}$. In this case, Bob must first search all of $\|blue\|_{Bob}$, then the ones he does not consider blue until he finds b there. His expected search effort can be equated to the cardinality of the set $\|blue\|_{Bob}$ plus half that of the complement of this set. In this "unlucky" scenario, Ann's utterance has led Bob astray: without information about the colour of the book, he could have expected to examine only half the domain.

In Parikh's story, Ann and Bob both used a crisp (i.e., non-vague) concept 'blue'. In what follows, we will argue that it would be advantageous for Bob (and, by extension, for Ann, who wants the book to be found) if Bob was able to rise

[1] Differences between speakers are particularly difficult to accommodate in *epistemicist* (i.e., "vagueness as ignorance") approaches to vagueness, which often assume that there is always only one true answer to the question "Is this person tall". See [7], Chapter 7.

above thinking in terms of a simple dichotomy between blue and non-blue. Bob might argue, for example, that if the target book is not found among the ones he considers blue, then it is most likely to be one that he considers borderline blue; so after inspecting the books he considers blue, he would be wise to inspect these borderline cases. He might even think of the books as arranged in order of their degree of blueness, and start earching the ones that are most typically blue, followed by the ones that are just slightly less blue, and so on.

Colours are complex, multi-dimensional things. For simplicity, we shall focus on the one-dimensional word tall. Thus, of any two extensions that the word may be assigned in a given situation, one must always be a subset of the other ($\|tall\|_{Ann}\| \subseteq \|tall\|_{Bob}$ or $\|tall\|_{Bob}\| \subseteq \|tall\|_{Ann}$, or both). More crucially, let us abandon the assumption that Ann and Bob must always think of the words in question as expressing a crisp dichotomy.[2] The story of the stolen diamond is set in Beijings Forbidden City, long ago:

> A diamond has been stolen from the Emperor and, security being tight in the palace, the thief must have been one of the Emperors 1000 eunuchs. A witness sees a suspicious character sneaking away. He tries to catch him but fails, getting fatally injured in the process. The scoundrel escapes. With his last breath, the witness reports "The thief is tall!", then gives up the ghost. How can the Emperor capitalize on these momentous last words? ([7], Chapter 9.)

Suppose the Emperor thinks of tall as a dichotomy, meaning taller than average, for instance. In this case, his men will gather all those eunuchs who are taller than average, perhaps about 500 of them. In the absence of any further clues, he should expect to search an average of as many as 250 tall people (i.e., half of the total number). Matters get worse if the witness has used a more relaxed notion of tall than the Emperor. If this mismatch arises, it is possible that the perpetrator will not be among the eunuchs whom the Emperor considers to be tall. Since the Emperor's concept "tall" makes no distinctions between people who are not tall, the Emperors men can only search them in arbitrary order. In other words, he first searches 500 eunuchs in vain, then an expected 0.5 * 500 = 250, totalling 750. Analogous to the previous section, the Emperor would have been better off without any description of the thiefs height, in which case he should have expected to search 0.5 * 1000 = 500 eunuchs. The Emperor could have diminished the likelihood of a false start by counting more eunuchs as tall. But in doing so, he would have increased the search times that are necessary to inspect all the eunuchs he considers tall. The only way to avoid the possibility of a false start altogether is by counting *all* eunuchs as tall, which would rob the witness statement of its usefulness.

[2] Parikh hints briefly at a related possibility in a footnote, without discussing its implications: "It may be worth pointing out that probably Bob does have another larger set of Bluish(Bob) books which includes both Blue(Ann) and Blue(Bob). After looking through Blue(Bob), he will most likely look only through the remaining Bluish(Bob) books." ([13], p. 533) See also our section 4, where expressions like "somewhat tall" are discussed.

If the Emperor thinks of tall as vague, however, then he might separate the eunuchs into three groups: the ones who are definitely tall, the ones who are definitely not tall, and the borderline cases characteristic of vague concepts. For concreteness, assume 100 eunuchs are definitely tall, 500 are definitely not tall, and 400 are doubtful. Surely, the eunuchs in the "definitely tall" category are more likely to be called tall than the ones in the "doubtfull" category, while no one in the "definitely not tall" category could be called tall. To put some figures to it, let the chance of finding the thief in the group of 100 be 50% and the chance of finding him in the doubtful group of 400 likewise. Under this scenario, it pays off to search the "definitely tall" eunuchs first, as one may easily verify. In other words, the Emperor benefits from regarding tall as containing borderline cases (i.e., being vague). This thought experiment suggests that borderline cases, and hence vagueness, can facilitate search, because borderline cases allow us to distinguish more finely than would be possible if all our concepts were dichotomies. If your language contains only dichotomous concepts then separating the eunuchs into three different groups does not make sense: there are tall eunuchs, not-tall ones, and thats it. But if you understand tall to have borderline cases then you can distinguish between the different people whom you do not consider tall, as well as between the ones you consider tall and all the others.

But if distinguishing between three different categories is better than distinguishing between just two, then it might be even better to distinguish even more finely. The Emperor can do even better than was suggested above if he uses a *ranking* strategy. Suppose he has the eunuchs arranged according to their heights. First the tallest eunuch is searched, then the tallest but one, and so on, until the diamond is found. This strategy is faster than each of the other ones if we assume that *the taller a person is, the more likely the witness is to have described him as tall.* Under this assumption, the same type of advantage obtains as in the previous case (where only borderline cases were acknowledged), but at a larger scale. – Note that we are ascribing a ranking strategy to the Emperor (i.e., the hearer) only. For all we know, the witness may be ignorant of the eunuchs heights while only possessing a rough impression of that of the thief. The Emperor and his men, by contrast, can rank the eunuchs at their ease.

This argument suggests an interesting possible *rationale* for understanding 'tall' as involving borderline cases or degrees, namely that this allows a more efficient search than would have been possible under a dichotomous understanding of these words. But borderline cases and degrees are the hallmark of vagueness. Consequently, if the argument is correct then we have found a so-far unnoticed *rationale* for vagueness: vagueness can facilitate search.

3 Towards a formal development of the argument

We aim to show that, given a dichotomous model, it is normally possible to define a closely resembling vague model, which has a higher utility than the original crisp one, where utility is formalised by the amount of search that has to be undertaken by a hearer who uses the model in question. Let's assume that

$\mathcal{A}$ is a standard two-valued model of the word 'tall' as defined on a domain D of people, where some people in D are tall (such people are in the extension $\|tall\|$) and others are not (such people are in the extension $\|\overline{tall}\|_{\mathcal{A}}$). $\mathcal{B}$, by contrast, has a truth-value gap: according to $\mathcal{B}$, there are not only tall and not-tall people, but borderline cases as well (such people are in the extension $\|?tall?\|_{\mathcal{B}}$). As before, the search effort implied by a model $\mathcal{X}$ (abbreviated $s(\mathcal{X})$) will be formalised as the *expected* number of elements of D that the hearer will have to examine, under the assumption that she goes on searching until the intended referent (i.e., the man with the diamond in his pocket) is found. For simplicity, assume that the models $\mathcal{A}$ and $\mathcal{B}$ call exactly the same people tall, so both assign the same extension to $\|tall\|$.[3]

3.1 The advantage of allowing borderline cases

Let us compare the models $\mathcal{A}$ and B above. Focussing on the witness' reference to the thief (t), there are three different types of situations. (In what follows, $card(X)$ abbreviates "the cardinality of X").

> **Type 1.** $t \in \|tall\|$. In this case, $s(\mathcal{A}){=}s(\mathcal{B})$, because the same sets are searched in both cases.
> **Type 2.** $t \in \|?tall?\|_{\mathcal{B}}$. In this case, $s(\mathcal{A}){>}s(\mathcal{B})$, so B leads to a lower search effort than $\mathcal{A}$. In other words, the model with borderline cases (i.e., model $\mathcal{B}$) incurs an advantage over the one that does not (i.e., $\mathcal{A}$). The size of the advantage is $1/2(card(\|\overline{tall}\|_{\mathcal{B}}))$.
> **Type 3.** $t \in \|\overline{tall}\|_{\mathcal{B}}$. In this case, $s(\mathcal{B}){>}s(\mathcal{A})$, in other words the model with borderline cases incurs a disadvantage. The size of the disadvantage is $1/2(card(\|?tall?\|_{\mathcal{B}}))$.

Proofs of these claims use standard reasoning about probability. Consider Type 2, for example, where the thief t is borderline tall. Given our assumptions, this implies $t \in \|\overline{tall}\|_{\mathcal{A}}$. We can measure the hearer's search effort implied by the model $\mathcal{A}$ as $s(\mathcal{A}) = card(\|tall\|) + 1/2(card(\|\overline{tall}\|_{\mathcal{A}}))$. The search effort implied by $\mathcal{B}$ is $s(\mathcal{A}) = card(\|tall\|) + 1/2(card(\|?tall?\|_{\mathcal{B}}))$, so $s(\mathcal{A}){>}s(\mathcal{B})$ if $card(\|\overline{tall}\|_{\mathcal{A}}) > card(\|?tall?\|_{\mathcal{B}})$, which is true given that (as we assumed) $\|\overline{tall}\|_{\mathcal{B}} \neq \emptyset$. The size of the advantage is $1/2(card(\|\overline{tall}\|_{\mathcal{A}})) - 1/2(card(\|?tall?\|_{\mathcal{B}}))$, which equals $1/2(card(\|\overline{tall}\|_{\mathcal{B}}))$.

What we really like to know is the expected search effort *a priori*, when it is not known in which of the three Types of situations (listed above) we are (i.e., whether the thief is in $\|tall\|$, in $\|?tall?\|_{\mathcal{B}}$, or in $\|\overline{tall}\|_{\mathcal{B}}$). Let "tall(x)" (in double quotes) say that the witness calls x tall, then the following hypothesis, H, seems highly plausible:

> **Hypothesis** H :
> $\forall xy((x \in \|?tall?\|_{\mathcal{B}} \wedge y \in \|\overline{tall}\|_{\mathcal{B}}) \rightarrow p(\text{"tall(x)"}) > p(\text{"tall(y)"}))$.

[3] Other assumptions can have similar consequences. See e.g. section 2, where we assumed that $\|\overline{tall}\|_{\mathcal{A}} = \|\overline{tall}\|_{\mathcal{B}}$.

(Note that H is not dependent on the size of $\|?tall?\|_{\mathcal{B}}$ and $\|\overline{tall}\|_{\mathcal{B}}$.) In justification of H: a person in $\|\overline{tall}\|_{\mathcal{B}}$ is considered clearly not-tall by the emperor so, although it cannot be ruled out that the witness described the same person as tall, such mismatches cannot occur too frequently amongst people who speak the same language, or else communication will break down. It would be far less unusual to see a person in the borderline area $\|?tall?\|_{\mathcal{B}}$ being described as tall: intuitively speaking, this category exists precisely to take account of the fact that the individuals in it may be considered tall by some but not all speakers. Hypothesis H has important consequences because, given that only one individual is called "tall" and this individual is the (only) thief, it follows that

$$\forall xy((x \in \|?tall?\|_{\mathcal{B}} \wedge y \in \|\overline{tall}\|_{\mathcal{B}}) \rightarrow p(thief(x)) > p(thief(y))).$$

Consequently, it is advantageous to search (all of) $\|?tall?\|_{\mathcal{B}}$ before $\|\overline{tall}\|_{\mathcal{B}}$. It follows that $s(\mathcal{A}) > s(\mathcal{B})$. In other words: given a dichotomous model, it is always possible to find a non-dichotomous model (i.e., with borderline cases) which agrees with it on all positive cases and which implies a smaller search effort on the part of the hearer.

3.2 The advantage of degrees and ranking

To develop a formal take on what happens when a concept like "tall" is seen as having *degrees*, let us contemplate a degree model $\mathcal{C}$, alongside the dichotomous model $\mathcal{A}$ and the three-valued model $\mathcal{B}$. Without loss of generality we can assume that $\mathcal{C}$ assigns real-valued truth values in $[0,1]$ to each person in D. As is customary in Fuzzy Logic ([14], [9]), among other systems, let $\mathcal{C}$ assign the value 0 to the shortest person and 1 to the tallest, while taller people are be assigned values that are not lower than those assigned to shorter ones.

In the present context, the crucial advantage of degree models over 2- or 3-valued ones is that degree models tend to make finer distinctions. 2-valued models (i.e., dichotomous ones) are able to distinguish between two kinds of people (the tall ones and the not-tall ones), and 3-valued models (i.e., ones with a truth-value gap) are able to distinguish between three. Degree models have the capacity to distinguish between many more people – if need be, a mathematical continuum of them. Where this happens, the advantages are analogous to the previous subsection.

Suppose, for example, that the domain contains ten individuals: $a1$, $a2$, $b1$, $b2$, $c1$, $c2$, $d1$, $d2$, $e1$, and $e2$, where $a1$ and $a2$ have (approximately) the same height, so do $b1$ and $b2$, and so on. (The number of members of each of the types a-e is immaterial: instead of all having two members, each of them could have any positive number of members.) Assume that the Emperor assigns "fuzzy" truth values as follows:

$v(Tall(a1)) = v(Tall(a2)) = 0.9,$
$v(Tall(b1)) = v(Tall(b2)) = 0.7,$
$v(Tall(c1)) = v(Tall(c2)) = 0.5,$
$v(Tall(d1)) = v(Tall(d2)) = 0.3,$
$v(Tall(e1)) = v(Tall(e2)) = 0.1.$

Recall that the witness described the thief as "tall". It is not farfetched to think that $a1$ and $a2$ are more likely targets of this description than $b1$ and $b2$, while these two are more likely targets than $c1$ and $c2$, and so on. The Emperor should therefore start looking for the diamond in the pockets of the two tallest individuals, then in those of the two next tallest ones, and so on. The idea is the same as in the previous subsection, except with five rather three levels of height: under the assumptions that were made, this search strategy is quicker than the previous two.

This example suggests that the key to the success of this strategy is the ability to *rank* the indivuals in terms of their heights, assuming that this corresponds to a ranking of their likelihood of being called "tall". Whenever this ability results in finer distinctions than 2- or 3-valued models, degree models lead to diminished search effort. To see how this works, it suffices to realise that the hypothesis H', a minor variant of the earlier hypothesis H, is highly plausible:

Hypothesis $H' : \forall xy(v(tall(x)) > v(tall(y)) \rightarrow p(\text{“tall(x)”}) > p(\text{“tall(y)”}))$.

Given the correlation between $v(tall(x)$ and the height of x, hypothesis H' says that taller individuals have a higher likelihood of being called tall than smaller ones. It follows from this hypothesis that it is advantageous to start searching the tallest individual (or individuals) in the domain, then the next tallest, and so on. If the domain contains individuals of four or more height levels (i.e., at least four different truth values of the form $v(Tall(x))$ then this leads to an expected search time associated with the degree model $\mathcal{C}$ which is smaller than that associated with model $\mathcal{B}$, which has a truth-value gap. So, given a three-valued model, it is always possible to find a degree model that respects the distinctions made by the three-valued model and that implies an even smaller search effort on the part of the hearer.

As it stands, hypothesis H' might be false. To see why, suppose the height of x is 210cm, while that of y is 190cm. It could be argued that, in this situation, y is *more* (instead of less) likely to be called tall than x, because x is quite untypical for someone designated as tall: x might be more likely to be called "extremely tall", or even a "giant". Wrinkles of this kind can only be ironed out by empirical research, which should tell us individuals of what height are most likely to be called "tall", "extremely tall", and so on. The outcome of these empirical investigations should then lead to a modified version of H', which will tell the Emperor who to search first, based on the heights of the individuals in question. Like the original H', the modified hypothesis would allow the Emperor to benefit from a degree model.

4 Discussion

We have argued that quasi-continuous domains make it difficult to align the meanings of the predicates defined over them: there are bound to be things that one person calls 'large' (or 'blue', or 'warm') that another person does not. Given such mismatches – which do not exist in standard game-theoretical analyses of vagueness – we have argued that it is to the hearer's advantage to distinguish shades of meaning in a way that is typical for vague concepts, namely using borderline cases or degrees. This argument suggests an answer to the *pragmatic* question that we asked in our Introduction which differs notably from the ones offered in the literature so far (see [6], [7]). To the extent that it supports degree-based models, ranging from Fuzzy Logic or probabilistic logic (e.g., [15], [16]) to Kennedy-style 2-valued semantics [10], our analysis also appears to shed light on the *logical-semantic* questions surrounding vagueness. – Let us discuss some possible objections against our argument.

Objection 1. It might be argued that the benefits that we ascribed to 3-valued and many-valued models can, in fact, also be obtained from 2-valued models. According to this view, the user of a 2-valued model is just as able to make fine distinctions as the user of any other kind of model. One can imagine a semantic and a syntactic version of this argument. The semantic version would argue that an intelligent user of a 2-valued model should be aware that *other* (2-valued) models may exist. She could argue, for example, that taller people are counted as tall by *more* models than less tall people. Clearly then, it pays to start searching amongst those people who are counted as tall by the largest set of models, that is, amongst the tallest people. The *syntactic* version of this argument would say that a person who is "quite" tall is a more likely to be called tall than someone is "somewhat" tall, who is more likely to be called tall than someone who is "a little bit on the tall side perhaps", and so on; therefore, after unsuccessfully searching all the people who are downright tall, the hearer should direct her attention to the people who are quite tall, somewhat tall, and so on.

But all objections of this kind presuppose that Bob's understanding of "large" goes beyond a simple dichotomous model. A language user who reasons as in the semantic version of this objection, for instance, knows that "tall" can have many different thresholds (corresponding with the different models), and reasons about these different thresholds. Essentially, this amounts to a *supervaluational* account of vagueness (e.g. [17]). The counterargument against the syntactic version of the objection is analogous: going beyond what the witness said, by exploring the extension of qualifications like "somewhat tall", does not make sense unless one is aware that the word "tall" is used differently by different people. Once again, if the Emperor followed this strategy, we would be justified in ascribing to him an understanding of "tall" as a vague concept.

Objection 2. Why did the witness keep us guessing, by using a vague concept? Why did he NOT say "the thief is 185cm tall", or something precise like that? – It is true that the utterance, "the thief is 185cm tall" might have been more helpful, but it is most naturally understood as vague too. To see this, note that the speaker may not know the exact height of the thief (nor his rank in terms of height), since he may have only a rough impression of his height, while he might know even less about other people's heights. Consequently, the speaker is not in a position to pass on the exact height of the thief. For this reason, an utterance like "the thief is 185cm tall" would tend to be interpreted as true of a person who is, for example, 184.4cm. At what height exactly the assessment starts being false would be difficult to say. Its meaning is perhaps best captured by a Gaussian function that asserts that 185cm is the most likely height, with other heights becoming less and less likely as they are further removed from 185cm. If such a vague estimate of the thief's height comes more naturally to human speakers than a precise assessment (e.g., "the thief's height is 185cm plus or minus 2cm") then Lipman's question can be repeated: why is this the case? Why, in other words, do statements in which speakers estimate heights tend to be vague? This new question can be answered in the same way as the question on which we focussed in this paper, by pointing out that a crisp concept like "height = 185cm plus or minus 2cm" would suffer from the same lack of flexibility as a crisp concept of "tall". Like before, vagueness allows speakers to deal flexibly with the differences among each other. Bob should start his search by focussing on individuals very close to 185cm, fanning out in both directions (i.e., below and above 185cm) until he has found the culprit.

Our account suggests a somewhat heritical view of *reference*. It is, of course, commonly understood that adjectives like "tall" are not intersective, but we can take this idea a step further. When a speaker refers to someone as "the tall man with the diamond in his pocket", one might believe that the hearer should consider the set of all men, intersect this with the set of individuals that have diamonds in their pockets, then intersects the result with the set of all people above a certain, contextually determined, height. The reason why this contextualised intersective interpretation will not work is *not* just that the hearer has incomplete information about the height standards employed by the speaker. The speaker did not necessarily employ any particular height standard; rather, she suggested that the best way to find the (unique) man with the diamond in his pocket is to start searching all the men in order of their height, because this is the quickest way to find the one with the diamond in his pocket.

Objection 3. It can be argued that a 3-valued model such as $\mathcal{A}$ falls short of making "tall" a vague concept, given that its boundaries (i.e., between $\|tall\|$, $\|?tall?\|$, and $\|\overline{tall}\|$) are crisp instead of vague. One might even go further and argue that the same is true for the many-valued models discussed in section 3.2, since these, too, assign definite truth values to each statement of the form "Tall(x)". I would counter that, if these models are seen as failing to model genuine (i.e., higher-order) vagueness, then it is difficult to see what models *do* model genuine vaguenes. Certainly very few of the models on the theoretical market (see e.g. [1]) go further than many-valued models in acknowledging vagueness. Essentially, in this paper, I have taken the pragmatic question about vagueness to be "Why does language not make do with simple dichotomous concepts?"

Objection 4. Lipman, in [4], proves a game-theoretical theorem (framed within a standard model as proposed in Crawford and Sobel 1982) stating that, given a vague predicate P, there must always exist a non-vague predicate P' where the utility of P' is at least as great as that of P. It might be thought that this contradicts the main claim of the present paper, but this is not the case. To prove his theorem, Lipman makes various assumptions which our analysis does not share. One of these assumptions is that a vague predicate is a *probability distribution* over functions that assign messages to heights. This is known as a mixed strategy, as opposed to a pure strategy, which is just a function from heights to messages. We have adopted a different attitude towards vagueness, without probability distributions. A second, and even more crucial assumption on which Lipman's theorem rests is that there are no mismatches between speaker and hearer. In particular, when a pure strategy is adopted by the speaker, he assumes that the hearer knows what this strategy is. Our own investigations, of course, start from a very different assumption, for which there exists ample empirical evidence (e.g. [11], [13], [12]) namely that mismatches between speakers' and hearers' understanding of concepts like 'tall' are unavoidable (i.e., perfect alignment would be a miracle).

References

1. Keefe,R. and Smith,P. *Vagueness: a Reader.* MIT Press, Cambridge Mass. and London. (1997)
2. van Rooij, R. Vagueness and Linguistics. To appear in *Handbook of Vagueness.*
3. Lipman, B. Economics and Language. In the "Comments" section of [5]. (2000)
4. Lipman, B. Why is language vague? Working paper, Dept. of Economics, Boston University. (Version of Dec. 2006.)
5. Rubinstein, A. *Economics and Language: Five Essays.* Cambridge: Cambridge University Press. (2000)
6. van Deemter, K. Utility and Language Generation: The Case of Vagueness. *Journal of Philosophical Logic* **38** (6), pp.607-632. (2009)
7. van Deemter, K. *Not Exactly: in Praise of Vagueness.* Oxford University Press (Jan. 2010). Web site associated with the book: `http://www.csd.abdn.ac.uk/∼ kvdeemte/NotExactly`.
8. Crawford, V. and Sobel, J. Strategic information transmission. *Econometrica* **50**, pp. 1431-1451. (1981)
9. Bergmann, M. *An introduction to many-valued and fuzzy logic: semantics, algebras, and derivation systems.* Cambridge University Press. (2008)
10. Kennedy, C. Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives. Linguistics and Philosophy **30** (1). (2007)
11. Hilbert, D.R. Color and Color Perception: A Study in Anthropocentric Realism. CSLI Lecture Notes 9. Stanford, Calif.: CSLI Publications. (1987)
12. Reiter, E., Sripada, J., Hunter, J., Yu, J., and Davy, I. Choosing words in computer-generated weather forecasts. *Artificial Intelligence* **167**, pp.137-96. (2005)
13. Parikh, R. Vagueness and Utility: the semantics of common nouns. *Ling. and Phil.* **17**: pp.521-535. (1994)
14. Zadeh, L. Fuzzy Sets. *Information and Control* **8** (3).: pp.338-53. (1965)
15. Edgington, D. Validity, uncertainty and vagueness, Analysis 52: pp. 193204. (1992)
16. Edgington, D. Vagueness by degrees. (1996) In [1].
17. Fine, K. Vagueness, truth and logic, Synthese 30: 265300. (1975)